RAW SEQUENCE LISTING DATE: 06/12/2001
PATENT APPLICATION: US/09/498,556 TIME: 15:06:48

Input Set : N:\Crf3\RULE60\09498556.txt
Output Set: N:\CRF3\06122001\I498556.raw

ENTERED

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Vlasuk, George Phillip
Stanssens, Patrick Eric Hugo
Meesens, Doris H.lda Lieven
Lauwereys, Marc Josef
Laroche, Yves Rene
Jespers, Laurent Stephane
Ganseman, Yannick Georges Jozef
Hoyle, Matthew
Bergum, Peter W.

(ii) TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE INHIBITORS AND ANTICOAGULANT PROTEIN

(iii) NUMBER OF SEQUENCES: 256

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Lyon & Lyon
(B) STREET: 633 West Fifth Street
Suite 4700
(C) CITY: Los Angeles
(D) STATE: California
(E) COUNTRY: U.S.A.
(F) ZIP: 90071

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: IBM P.C. DOS 5.0
(D) SOFTWARE: Word Perfect 5.1

(vi) CURRENT APPLICATION DATA:
C--> (A) APPLICATION NUMBER: US/09/498,556
C--> (B) FILING DATE: 04-Feb-2000

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/639,456
(B) FILING DATE: April 17, 1997
(A) APPLICATION NUMBER: PCT/US95/13231
(B) FILING DATE: October 17, 1995
(A) APPLICATION NUMBER: US 08/486,399
(B) FILING DATE: June 5, 1995
(A) APPLICATION NUMBER: US 08/486,397
(B) FILING DATE: June 5, 1995
(A) APPLICATION NUMBER: US 08/486,380
(B) FILING DATE: June 5, 1995
(A) APPLICATION NUMBER: US 08/461,965
(B) FILING DATE: June 5, 1995
(A) APPLICATION NUMBER: US/326,110
(B) FILING DATE: October 18, 1994

RAW SEQUENCE LISTING DATE: 06/12/2001
PATENT APPLICATION: US/09/498,556 TIME: 15:06:48

Input Set : N:\Crf3\RULE60\09498556.txt
Output Set : N:\CRF3\06122001\I498556.raw

```
80    (viii) ATTORNEY/AGENT INFORMATION:
82           (A) NAME: BIGGS, J'ANNE L.
84           (B) REGISTRATION NUMBER: 30,158
86           (C) REFERENCE/DOCKET NUMBER: 216/270
87      (ix) TELECOMMUNICATION INFORMATION:
89           (A) TELEPHONE: (213) 489-1600
90           (B) TELEFAX: (213) 955-0440
91           (C) TELEX: 67-3510
93  (2) INFORMATION FOR SEQ ID NO: 1:
97       (i) SEQUENCE CHARACTERISTICS:
99           (A) LENGTH: 231 base pairs
100          (B) TYPE: nucleic acid
101          (C) STRANDEDNESS: single
102          (D) TOPOLOGY: linear
104      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
106  AAGGCATACC CGGAGTGTGG TGAGAATGAA TGGCTCGACG ACTGTGGAAC TCAGAAGCCA      60
107  TGCGAGGCCA AGTCAATGGA GGAAGTCCTT GAGGAGGAAG ATCCGATATG CCGCTCACGT     120
108  GGTTGTTTAT TATTCTCTGT TTGCGTATGC AAAGACGGAT TCTACAGAGA CACGGTGATC     180
109  GGCGACTGTG TTAAGAAAGA AGAATCCGAC CAACATGAGA TTATACATGT CTGA           234
113 (2) INFORMATION FOR SEQ ID NO: 2:
115      (i) SEQUENCE CHARACTERISTICS:
117          (A) LENGTH: 234 base pairs
118          (B) TYPE: nucleic acid
119          (C) STRANDEDNESS: single
120          (D) TOPOLOGY: linear
122      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
124  AAGGCATACC CGGAGTGTGG TGAGAATGAA TGGCTCGACG TCTGTGGAAC TAAGAAGCCA      60
126  TGCGAGGCCA AGTCAATGGA GGAAGAGGAG GAAGATCCGA TATGCCGATC ATTTTCTTGT     120
128  CCGGGTCCCG TCGTTCTGGT ATGCTAAGAC GGATTCTACA GAGACACGGT GATCGGCGAC     180
129  TGTGTTAAGG AAGAAGAATC CGACTAACAT GAGATTATAC ATGTCTGA                  228
134 (2) INFORMATION FOR SEQ ID NO: 3:
136      (i) SEQUENCE CHARACTERISTICS:
138          (A) LENGTH: 461 base pairs
139          (B) TYPE: nucleic acid
140          (C) STRANDEDNESS: single
141          (D) TOPOLOGY: linear
143      (vi) ORIGINAL SOURCE:
145          (A) ORGANISM: Amycolatopsis caninum
147      (ix) FEATURE:
149          (A) NAME/KEY: Coding Sequence
150          (B) LOCATION: 2....371
153      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
155  GAATTCCGCT ACTACTTAAC A ATG AAG ATG CTT TAC GCT ATC GCT ATA ATG       51
156                         Met Lys Met Leu Tyr Ala Ile Ala Ile Met
157                          1               5                  10
159  TTT CTC CTG GTA TCA TTA TGC AGC GCA AGA ACA GTG AGG AAG GCA TAC       99
160  Phe Leu Leu Val Ser Leu Cys Ser Ala Arg Thr Val Arg Lys Ala Tyr
161                  15                  20                  25
163  CCG GAG TGT GCT GAG AAT GAA TGG CTC GAC GAC TGT GGA ACT CAG AAG      147
```

RAW SEQUENCE LISTING DATE: 06/12/2001
PATENT APPLICATION: US/09/498,556 TIME: 15:06:48

Input Set : N:\Crf3\RULE60\09498556.txt
Output Set: N:\CRF3\06122001\I498556.raw

```
164  Pro Glu Cys Gly Glu Asn Glu Trp Leu Asp Asp Cys Gly Thr Gln Lys
165                30                  35                  40
167  CCA TGC GAG GGT AAG TGC AAT GAG GAA CCT CCT GAG GAG GAA GAT CCG      195
168  Pro Cys Glu Ala Lys Cys Asn Glu Glu Pro Pro Glu Glu Glu Asp Pro
169           45                  50                  55
171  ATA TGC CGC TCA CGT GGT TGT TTA TTA CCT CCT GCT TGC GTA TGC AAA      243
172  Ile Cys Arg Ser Arg Gly Cys Leu Leu Pro Pro Ala Cys Val Cys Lys
173       60                  65                  70
175  GAC GGA TTC TAC AGA GAC ACG GTG ATC GGC GAC TGT GTT AGG GAA GAA      291
176  Asp Gly Phe Tyr Arg Asp Thr Val Ile Gly Asp Cys Val Arg Glu Glu
177  75                  80                  85                  90
179  GAA TGC GAC GAA CAT GAG ATT ATA CAT GTC T GAACGAGAAA GCAACAATAA CC    344
180  Glu Cys Asp Glu His Glu Ile Ile His Val
181                  95                 100
183  AAAGGTTCCA ACTCTCGCTC TGCAAAATCG CTAGTIGGAT GTCTCTTTTG CGTCCGAATA     404
185  GTTTTAGTTG ATGTTAAGTA AGAATCCCTG CTGGAGAGAA TAAAGCTTTC CAACTCC        461
187  (2) INFORMATION FOR SEQ ID NO: 4:
189       (i) SEQUENCE CHARACTERISTICS:
191            (A) LENGTH: 77 amino acids
192            (B) TYPE: amino acid
193            (D) TOPOLOGY: linear
195       (ii) MOLECULE TYPE: peptide
197       (vi) ORIGINAL SOURCE:
199            (A) ORGANISM: Ancylostoma caninum
201       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
203  Lys Ala Tyr Pro Glu Cys Gly Glu Asn Glu Trp Leu Asp Asp
204   1               5                  10
206  Cys Gly Thr Gln Lys Pro Cys Glu Ala Lys Cys Asn Glu Glu
207   15                  20                  25
209  Pro Pro Gln Gln Glu Asp Pro Ile Cys Arg Ser Arg Gly Cys
210        30                  35                  40
212  Leu Leu Pro Pro Ala Cys Val Cys Lys Asp Gly Phe Tyr Arg
213           45                  50                  55
215  Asp Thr Val Ile Gly Asp Cys Val Arg Glu Glu Glu Cys Asp
216                60                  65                  70
218  Gln His Glu Ile Ile His Val
219                75
222  (2) INFORMATION FOR SEQ ID NO: 5:
224       (i) SEQUENCE CHARACTERISTICS:
226            (A) LENGTH: 455 base pairs
227            (B) TYPE: nucleic acid
228            (C) STRANDEDNESS: single
229            (D) TOPOLOGY: linear
231       (vi) ORIGINAL SOURCE:
233            (A) ORGANISM: Ancylostoma caninum
235       (ix) FEATURE:
237            (A) NAME KEY: Coding Sequence
238            (B) LOCATION: 22...515
240       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
```

RAW SEQUENCE LISTING DATE: 06/12/2001
PATENT APPLICATION: US/09/498,556 TIME: 14:06:48

Input Set : N:\Crf3\RULE60\09498556.txt
Output Set: N:\CRF3\06122001\I498556.raw

```
244  GAATTCCGCT ACTACTCAAC A ATG AAG ATG CTT TAC GCT ATC GCT ATA AAG       51
245                          Met Lys Met Leu Tyr Ala Ile Ala Ile Met
246                           1               5                  10
248  TTT CTC CTG GTG TCA TTA TGC AGC ACA AGA ACA GTG AGG AAG GCA TAC       99
249  Phe Leu Leu Val Ser Leu Cys Ser Thr Arg Thr Val Arg Lys Ala Tyr
250                  15                  20                  25
252  CCG GAG TGT GGT GAG AAT GAA TGG CTC GAC GTC TGT GGA ACT AAG AAG      147
253  Pro Glu Cys Gly Glu Asn Glu Trp Leu Asp Val Cys Gly Thr Lys Lys
254              30                  35                  40
256  CCA TGC GAG GCC AAG TGC AGT GAG GAA GAG GAG GAA GAT CCG ATA TGC      195
257  Pro Cys Glu Ala Lys Cys Ser Glu Glu Glu Glu Glu Asp Pro Ile Cys
258          45                  50                  55
260  CGA TCA TTT TCT TGT CCG GGT CCC GCT GCT TGC GTA TGC GAA GAC GGA      243
261  Arg Ser Phe Ser Cys Pro Gly Pro Ala Ala Cys Val Cys Glu Asp Gly
262      60                  65                  70
264  TTC TAC AGA GAC ACG GTG ATC GGC GAC TGT GTT AAG GAA GAA GAA TGC      291
265  Phe Tyr Arg Asp Thr Val Ile Gly Asp Cys Val Lys Glu Glu Glu Cys
266  75                  80                  85                  90
268  GAC CAA CAT GAG ATT ATT CAT GTC TGAACGAGAG AGCAGTAATA ACCAAAGTTT C   346
269  Asp Gln His Glu Ile Ile His Val
270                  95
272  CAACTTTCGC TTTACAAAAT CGCTAGTTGG ATTTCTCCTT TGCGTGCGAA TAGTTTTAGT    406
273  TGATATTAAG TAAAAACCTCC TGTTGAAGAG AATAAAGCTT TCCAACTTT              455
275  (2) INFORMATION FOR SEQ ID NO: 6:
277       (i) SEQUENCE CHARACTERISTICS:
279            (A) LENGTH: 75 amino acids
280            (B) TYPE: amino acid
281            (D) TOPOLOGY: linear
283       (ii) MOLECULE TYPE: peptide
285       (vi) ORIGINAL SOURCE:
287            (A) ORGANISM: Ancylostoma caninum
289       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
291  Lys Ala Tyr Pro Glu Cys Gly Glu Asn Glu Trp Leu Asp Val Cys Gly
292   1               5                  10                  15
294  Thr Lys Lys Pro Cys Glu Ala Lys Cys Ser Glu Glu Glu Glu Glu Asp
295              20                  25                  30
297  Pro Ile Cys Arg Ser Phe Ser Cys Pro Gly Pro Ala Ala Cys Val Cys
298          35                  40                  45
300  Glu Asp Gly Phe Tyr Arg Asp Thr Val Ile Gly Asp Cys Val Lys Glu
301      50                  55                  60
303  Glu Glu Cys Asp Gln His Glu Ile Ile His Val
304  65                  70                  75
306  (2) INFORMATION FOR SEQ ID NO: 7:
308       (i) SEQUENCE CHARACTERISTICS:
310            (A) LENGTH: 81 amino acids
311            (B) TYPE: amino acid
312            (D) TOPOLOGY: linear
314       (ii) MOLECULE TYPE: peptide
316       (vi) ORIGINAL SOURCE:
```

RAW SEQUENCE LISTING DATE: 06/12/2001
PATENT APPLICATION: US/09/498,556 TIME: 15:06:48

Input Set : N:\Crf3\RULE60\09498556.txt
Output Set: N:\CRF3\06122001\I498556.raw

```
(A) ORGANISM: Amblyomma caninum
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
Arg Thr Val Asn Lys Ala Tyr Pro Glu Cys Gly Glu Asn Glu Trp Leu
 1               5                  10                  15
Asp Asp Cys Gly Thr Gln Lys Pro Cys Glu Ala Lys Cys Asn Glu Glu
            20                  25                  30
Pro Pro Glu Glu Glu Asp Pro Ile Cys Arg Ser Arg Gly Cys Leu Leu
        35                  40                  45
Pro Pro Ala Cys Val Cys Lys Asp Gly Phe Tyr Arg Asp Thr Val Ile
    50                  55                  60
Gly Asp Cys Val Asp Glu Glu Glu Cys Asp Gln His Glu Ile Ile His
65                  70                  75                  80
Val
(2) INFORMATION FOR SEQ ID NO: 8:
     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 79 amino acids
          (B) TYPE: amino acid
          (C) TOPOLOGY: linear
     (ii) MOLECULE TYPE: peptide
     (vi) ORIGINAL SOURCE:
          (A) ORGANISM: Amblyomma caninum
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
Arg Thr Val Asn Lys Ala Tyr Pro Glu Cys Gly Glu Asn Glu Trp Leu
 1               5                  10                  15
Asp Val Cys Gly Thr Lys Lys Pro Cys Glu Ala Lys Cys Ser Glu Glu
            20                  25                  30
Glu Glu Glu Asp Pro Ile Cys Arg Ser Phe Ser Cys Pro Gly Pro Ala
        35                  40                  45
Ala Cys Val Cys Glu Asp Gly Phe Tyr Arg Asp Thr Val Ile Gly Asp
    50                  55                  60
Cys Val Lys Glu Glu Glu Cys Asp Gln His Glu Ile Ile His Val
65                  70                  75
(2) INFORMATION FOR SEQ ID NO: 9:
     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 711 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear
     (vi) ORIGINAL SOURCE:
          (A) ORGANISM: Amblyomma ceylanicum
    (ix) FEATURE:
          (A) NAME KEY: Coding Sequence
          (B) LOCATION: 21...690
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
GAATTCACTA TTATCGAACA  ATG GCG GTG CTT TAT TCA GTA GCA ATA GCG        50
                       Met Ala Val Leu Tyr Ser Val Ala Ile Ala
                        1               5                  10
TTA CTA CTG GTA TTA CAA TGC AGT GGG AAA CCG AAC AAT GTG ATG ACT        98
Leu Leu Leu Val Ser Gln Cys Ser Gly Lys Pro Asn Asn Val Met Thr
```

**VERIFICATION SUMMARY** DATE: 06/12/2001
PATENT APPLICATION: **US/09/498,556** TIME: 15:06:49

Input Set : **N:\Crf3\RULE60\09498556.txt**
Output Set: **N:\CRF3\06122001\I498556.raw**

L:40 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:50 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:2909 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:66
L:2932 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:67
L:2955 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:68
L:3154 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:78
L:3185 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:79
L:3210 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:80
L:3238 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:81
L:3260 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:82
L:3283 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:83
L:3712 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:84
L:3736 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:85
L:3762 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:86
L:3863 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:87
L:4451 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:118
L:4479 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:119
L:4163 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:129
L:4195 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:130
L:4204 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:131
L:4231 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:132
L:4254 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:133
L:4276 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:134
L:4279 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:135
L:4301 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:136
L:4304 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:137
L:4326 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:138
L:4329 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:139
L:4352 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:140
L:4355 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:141
L:4377 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:142
L:4380 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:143
L:4402 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:144
L:4405 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:145
L:4428 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:146
L:4471 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:148
L:4494 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:149
L:4517 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:150
L:4541 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:151
L:4564 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:152
L:4587 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:153
L:4611 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:154
L:4634 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:155
L:4657 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:156
L:4680 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:157
L:4703 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:158
L:4726 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:159
L:4749 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:160

VERIFICATION SUMMARY DATE: 06/12/2001
PATENT APPLICATION: US/09/498,556 TIME: 15:06:49

Input Set : N:\Crf3\RULE60\09498556.txt
Output Set: N:\CRF3\06122001\I498556.raw

L:4773 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:155
L:4797 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:156
L:4819 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:157
L:4841 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:158